

Quantifying the Utility-Privacy Tradeoff in Direct Load Control Programs

Roy Dong, *Student Member, IEEE*, Alvaro A. Cárdenas, *Member, IEEE*, Lillian J. Ratliff, *Student Member, IEEE*, Henrik Ohlsson, *Member, IEEE*, and S. Shankar Sastry, *Fellow, IEEE*

Abstract—The modernization of the electrical grid and the installation of smart meters come with many advantages to control and monitoring. However, in the wrong hands, the data might pose a privacy threat. In this paper, we consider the tradeoff between smart grid operations and the privacy of consumers. We analyze the tradeoff between smart grid operations and how often data is collected by considering a realistic direct-load control example using thermostatically controlled loads, and we give simulation results to show how its performance degrades as the sampling frequency decreases. Additionally, we introduce a new privacy metric, which we call inferential privacy. This privacy metric assumes a strong adversary model, and provides an upper bound on the adversary’s ability to infer a private parameter, independent of the algorithm he uses. Combining these two results allow us to directly consider the tradeoff between better load control and consumer privacy.

Index Terms—privacy, smart grid, direct load control, cyber-physical systems

I. INTRODUCTION

DATA collected by the smart grid enables a multitude of advantages to all parties, including better efficiency in energy distribution, more reliability, and transparency to electric utility customers in their energy consumption. Smart grid data, however, also raises the issue of data privacy. Energy usage data is collected at larger scales and at unprecedented levels of granularity. Monitoring energy consumption at high granularity can allow the inference of detailed information about consumers’ lives. This includes the times they eat, when they watch TV, and when they take a shower [1]. Such information is highly valuable and will be sought by many parties, including advertising companies [2], law enforcement [3], and criminals [4].

In response to these concerns, governments, researchers, and organizations are working on privacy standards and policies to guide advanced metering infrastructure (AMI) deployments. Researchers have considered the issue of data privacy in smart grid infrastructures, and have proposed novel mechanisms for protecting the collected data (encryption, access control, and cryptographic commitments) [5], [6], by anonymization and aggregation [7], [8], and by preventing inferences and re-identification from databases that allow queries from untrusted third parties (via differential privacy) [9].

R. Dong, L. J. Ratliff, H. Ohlsson, and S. S. Sastry are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA. e-mail: {roydong, ratliff1, ohlsson, sastry}@eecs.berkeley.edu.

A. A. Cárdenas is with the Department of Computer Science, University of Texas, Dallas, TX, USA. email: alvaro.cardenas@utdallas.edu.

While all these previous proposals have strong contributions, none of them has addressed the Fair Information Practice (FIP) principle of data minimization. The NISTIR 7628 [10] expresses the data minimization principle in the smart grid context as:

Limit the collection of data to only that necessary for Smart Grid operations, including planning and management, improving energy use and efficiency, account management, and billing.

This same principle is included in several smart grid privacy recommendations including those published by the North American Energy Standards Board [11], DOE [12], the Texas Legislature and Public Utility Commission [13], and the California Public Utilities Commission (CPUC) [14].

Electric utilities who want to follow these privacy recommendations do not have a sound reasoning principle to help them decide how much data is too little or too much. Our goal in this paper is to start discussing scientifically sound principles that can help determine how much data to collect in order to achieve a certain level of functionality of the grid, and how much privacy is granted to consumers under this data collection policy.

To successfully understand the utility-privacy tradeoff in these smart grid operations, we must quantify two things. First, we must model the tradeoff between how often data is collected and performance of smart grid operations. Second, we must understand the tradeoff between how much data is collected and the inference an adversary can make about a consumer’s private information.

In this paper, our goal is to analyze these tradeoffs in the context of electricity load shaping (demand response). To quantify how much data is needed for smart grid operations, we consider how the performance of proposed direct load control (DLC) mechanisms change as fewer and fewer measurements are received by the controller. To quantify the privacy risk in these mechanisms, we use recent results in nonintrusive load monitoring (NILM) to give guarantees on when NILM algorithms will not be able to infer the device usage of a consumer from observing the aggregate power consumption of a building, acting as a certificate of privacy for the consumer.

This paper extends ideas presented in our previous work [15], where we used model predictive control (MPC) methods on a highly stylized DLC model to reduce unexpected demand across time, and analyzed the likelihood of the demand exceeding certain bounds as a function of the sampling rate. In contrast, in this paper we consider a more realistic DLC

model and provide a formal definition of privacy that allows us to quantify the privacy of different sampling policies.

We note that our work is complementary to other privacy policies being researched. Our analysis helps determine how much data to collect and how often it should be collected. Once this is in place, encryption, anonymization and aggregation techniques can be employed in tandem.

The closest previous work to the ideas in this paper is the work of Sankar, Kar, Tandon, and Poor [16], which considers the utility and privacy tradeoffs from smart grid data. Their approach, however, focuses on the *quality* of collected data (by using quantization, and coding theory); while in this paper we focus on the *quantity* of collected data (or more precisely, the sampling interval). To the best of our knowledge, this is the first paper considering a concrete model of smart grid operations and the tradeoff between its performance and the privacy of its consumers.

The rest of the paper is organized as follows. In Section II, we outline a DLC model studied in recent literature and consider how different sampling policies affect its performance. In Section III, we define our metric for privacy, called *inferential privacy*, inspired by the recent literature in NILM. Inferential privacy provides a strong adversary model, and give guarantees on when this adversary can infer private information about the consumer. In this paper, we are interested in which private information, seemingly unrelated to DLC, can be inferred from the data collected by AMIs. Finally, in Section IV, we conclude and discuss directions for future work.

II. DIRECT LOAD CONTROL PERFORMANCE

In this section, we consider the effects of lower sampling rates on smart grid operations. In particular, we focus on a direct load control (DLC) application using thermostatically controlled loads (TCLs) to manage load imbalances. As mentioned in Section I, for this paper, we restrict our scope to consider data sampling policies. Other ways to alter the privacy of a consumer participating in an advanced metering infrastructure (AMI) include adding noise to data, modifying how data is aggregated, and the duration of data retention. Such investigations are outside the scope of this paper, but are an active topic of research [5], [6], [7], [8], [9], [16], [17].

DLC has been a promising future direction for the smart grid for a variety of reasons. By controlling loads which can be modified without much impact on consumer satisfaction, we can allay many costs by shifting loads from peak demand and compensating for real-time load imbalances. Additionally, as renewable energy penetration increases, the generation side of power is growing more uncertain and will require demand flexibility.

TCLs, which are often heating, ventilation, and air conditioning (HVAC) systems for buildings, are a promising avenue for the implementation of DLC policies [18], [19]. This is due to the fact that buildings have a thermal inertia and can, in essence, store energy. Power consumption can be deferred and shifted while resulting in an imperceptible change in temperature.

Additionally, such DLC policies are being deployed today. For example, Pacific Gas & Electric deployed the SmartAC

program in Spring 2007 [20]. Another provider of demand response (DR) services has recruited over 1.25 million residential customers in DLC programs, and has deployed over 5 million DLC devices in the United States. In California, they have successfully curtailed over 25 MW of power consumption since 2007 [21]. As these programs are being deployed on a large scale, it is important to consider the privacy aspects of these programs.

In this paper, we consider one recently proposed DLC program for concreteness. We note that our contribution is a general framework for numerically analyzing the sensitivity of these DLC programs to different information collection policies. We consider this research to be complementary to other research in how parameters affect system performance [22], [23].

Additionally, our development focuses on air conditioning for notational simplicity, but similar statements can be made for heaters.

A. Thermostatically controlled load model

The model outlined in this section closely mirrors the model presented in [24]. There are other TCL and DLC models in the literature, e.g. [25], [26], and our analysis can similarly be applied to these models as well.

Let \mathcal{I} denote the set of TCLs participating in a DLC program. We model each TCL $i \in \mathcal{I}$ as a discrete-time difference equation:

$$\theta_i(k+1) = a_i\theta_i(k) + (1-a_i)[\theta_{a,i}(k) - m_i(k)\theta_{g,i}] + \epsilon_i(k) \quad (1)$$

In the above equations, $\theta_i(k)$ is the internal temperature of TCL i at time k , $\theta_{a,i}$ is the ambient temperature around TCL i , m_i is the control signal of TCL i , and ϵ_i is a noise process. The term $a_i = \exp(-h/(R_iC_i))$, where h is the sampling period, R_i is the thermal resistance of TCL i , and C_i is the thermal capacitance of TCL i . The θ_g term is the temperature gain when a TCL is in the ON state, and $\theta_g = R_iP_{\text{trans},i}$, where $P_{\text{trans},i}$ is the energy transfer rate of TCL i . Let P_i denote the power consumed by TCL i when it is in the ON state.

The local control for TCL i is modeled by the variable m_i . We assume the local controller does a basic ON/OFF hysteresis control based on its setpoint and deadband. For a cooling TCL, this is defined as:

$$m_i(k+1) = \begin{cases} 0 & \text{if } \theta_i(k+1) < \theta_{\text{set},i} - \delta_i/2 \\ 1 & \text{if } \theta_i(k+1) > \theta_{\text{set},i} + \delta_i/2 \\ m_i(k) & \text{otherwise} \end{cases} \quad (2)$$

In these equations, $\theta_{\text{set},i}$ and δ_i are the temperature setpoint and deadband of TCL i , respectively. If $m_i(k) = 1$, then we say that TCL i is in the ON state at time k , and similarly $m_i(k) = 0$ means that i is in the OFF state at k .

B. Direct load control objective

We consider DLC policies that attempt to compensate for load imbalances and defer demands from peak times by switching TCLs between the ON state and the OFF state. The marginal cost of peak loads and unexpected load imbalances is

responsible for a large portion of the preventable costs in the electricity grid; for a more detailed treatment of the benefits and impact of a DLC policy which can shave demand, we refer the reader to [27].

Formally, we consider the load imbalance as an exogenous variable. In particular, the centralized DLC operator is given some desired power trajectory P_{des} for the TCLs. The goal of the operator is to minimize the root-mean-square (RMS) error between the actual power consumed by the TCLs and the signal P_{des} , i.e. it wishes to minimize $\|\sum_{i \in \mathcal{I}} P_i m_i - P_{\text{des}}\|_2$.

C. Direct load control capabilities

To achieve the DLC objective, we assume the DLC operator has the capability of telling TCLs to switch modes between ON and OFF. This has the effect of tightening the deadband for TCLs. More explicitly, if the DLC operator issues a command to a TCL to switch from OFF to ON, the TCL turns on its air conditioner earlier than it would have in the absence of a control command, which is essentially lowering the ceiling of the deadband.

Note that this control policy maintains customer satisfaction in the sense that the thermal variation inside the TCL will not increase as a result of this control policy. However, the control commands may cause TCLs to switch more frequently, which may have the effect of faster depreciation and degradation of the TCL; as a heuristic for minimizing this effect, the controller we consider preferentially issues control commands to TCLs whose local hysteresis controller is likely to switch modes soon.

We assume the controller has access to the parameters $(a_i, \theta_{a,i}, \theta_{g,i}, \theta_{\text{set},i}, \delta_i, P_i)$ for each TCL $i \in \mathcal{I}$. In other words, the controller knows the dynamics of each TCL. However, it is only able to observe the signals $(\theta_i(k), m_i(k))$ for certain values of k , determined by the privacy-aware sampling policy.

Note here that constant parameters are considered non-private, whereas temporal information about the state inside the building is considered private. To understand the degree of privacy in these sampling policies, we can refer to the privacy metric introduced in Section III. We also note that situations may arise where some of these constant parameters are also private, and would like to consider this case in future work.

D. Direct load controller

In this section, we outline a direct load control policy inspired by work in the recent literature [18], [24].

Our model of a direct load controller is as follows. First, the controller maintains an estimate of the thermal state of each TCL. At time $k + 1$, the estimator uses the observation if it is available. If no measurement is available, it evolves the estimates according to the dynamics with known parameters α , under the assumption that $\epsilon_i(k) = 0$.

These estimates are used to issue control commands. Our controller takes a binning approach, as seen in recent research [18], [24]. Each TCL is assigned to a bin based on its thermal state relative to its deadband, and whether or not it is in the ON or OFF state. More formally, let N_{bin} be an even number denoting the number of bins our controller uses.

For the ON states, we assign $N_{\text{bin}}/2$ evenly spaced bins, and similarly for the OFF states.

Based on its estimate of how many TCLs are in each bin, the controller issues a command to each bin, stating what fraction of the TCLs in each bin should switch states. More concretely, the controller switches TCLs at time k based on the mismatch between the estimated power consumed $\sum_{i \in \mathcal{I}} P_i \hat{m}_i(k)$ at time k and the desired power consumption $P_{\text{des}}(k)$ at time k .

At the level of an individual TCL, the TCL can calculate which bin it is in, based on its true state $(\theta_i(k), m_i(k))$ and its deadband. When a TCL receives a command c , it will switch states with probability c . Using a probability allows the centralized controller to issue commands without broadcasting individual TCL identities, and without explicit knowledge of which TCLs will switch. Additionally, a TCL can decide whether or not to switch entirely on its own, without coordination or communication with other members of its bin.

E. Direct load control simulations

We summarize how simulations were generated for the framework just outlined. For this simulation, we assume each TCL consumes $P_i = 2.5$ kW when on, and we consider a DLC operator in control of 1000 TCLs. Parameters for each TCL i are drawn independently, from distributions based on recent studies of a 250 m² home [24], [18]. The time step h was chosen to be $h = 1$ minute, and the number of bins $N_{\text{bin}} = 10$.

The ambient temperature $\theta_a = 32^\circ\text{C}$ for all TCLs¹, and the noise process $\epsilon_i(k)$ is independent across k and distributed according to a $N(0, 0.0005)$ distribution for each k .

In normal operation, assuming the uniform distribution across the deadband and the Bernoulli distribution across ON/OFF states, the expected number of TCLs in the ON state is $1000/2 = 500$. Since each device consumes 2.5 kW when on, that means the expected power consumption at any time k is $500 \cdot 2.5 \text{ kW} = 1.25 \text{ MW}$.

California Independent System Operator (CAISO) market signals are given in 5 minute intervals [24], [28], so for simulations, the signal P_{des} is independently drawn from a $U(1.25 \text{ MW} \cdot 0.7, 1.25 \text{ MW} \cdot 1.05)$ distribution². That is, $P_{\text{des}}(k)$ is uniformly drawn for $k \in \{0, 5, 10, \dots\}$. For other values of k , we take the linear interpolation.

One simulation of the aggregate power consumption of all the TCLs is shown in Figure 1. Comparing the top plot with the middle and bottom plots, we can see that a DLC policy can reduce the load imbalance even when the controller does not always receive measurements. However, small unforeseen temperature deviations can cause the controller's performance to degrade if enough measurements are not provided, as seen by comparing the middle and bottom plots.

Additionally, the thermal state of one TCL is shown in Figure 2 for the uncontrolled case, the case where $h = 1$ minute,

¹For these simulations, we assumed that the ambient temperature is constant for the time window under consideration.

²These values were chosen as reasonable values for which energy consumption could be compensated. From simulations, we find that a larger interval is more difficult to track, as expected.

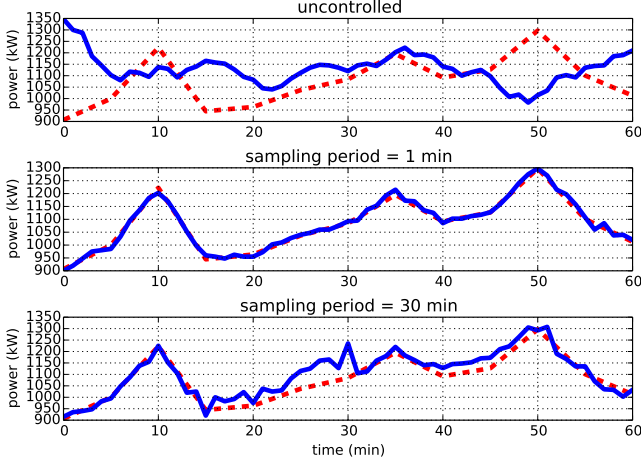


Fig. 1. A sample simulation of the aggregate power consumption of 1000 TCLs. The solid blue line represents the actual power consumption, and the dotted red line represents the desired power consumption. The top figure shows the power consumption in the absence of any control commands, the middle figure shows the power consumption with a sampling period of $h = 1$ minute, and the bottom figure shows the power consumption with a sampling period of $h = 30$ minutes.

and the case where $h = 30$ minutes. We can see that the temperature inside the TCL remains inside the deadband, resulting in no loss of comfort to the consumer, in all three cases.

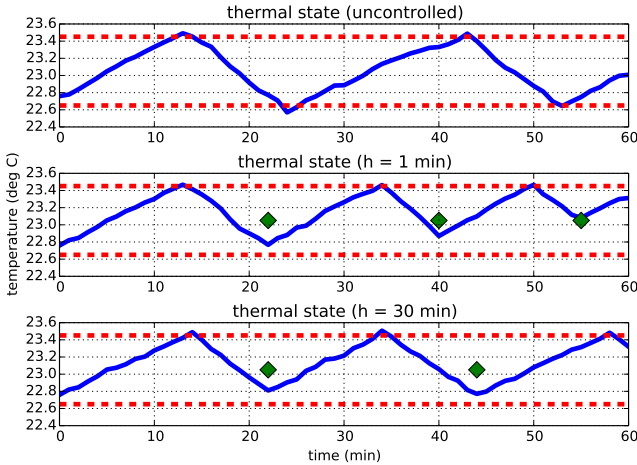


Fig. 2. The thermal state of one sample TCL. The top graph shows the thermal states of the TCL when there is no control. The middle and bottom graphs show the thermal states based on a controller that receives observations every $h = 1$ minute and $h = 30$ minutes, respectively. The dotted red lines indicate the deadband limits. The diamonds indicate when the DLC policy issued control commands to the TCL.

In Figure 3, we plot the error between the actual power consumption and the desired load imbalance compensation signal. First, we randomly drew a P_{des} signal and TCL parameters. Then, for this fixed P_{des} signal and TCL parameters, we ran 500 trials for each sampling period h , and we consider the empirical distribution of the difference between the actual

power consumed by all the TCLs and the desired power signal: $\sum_{i \in \mathcal{I}} P_i m_i - P_{\text{des}}$. We used the ℓ_1 norm on the error signal, so, if we assume a fixed price for spot market electricity purchases/sales throughout the hour interval, this is directly proportional to the cost the utility company must pay.

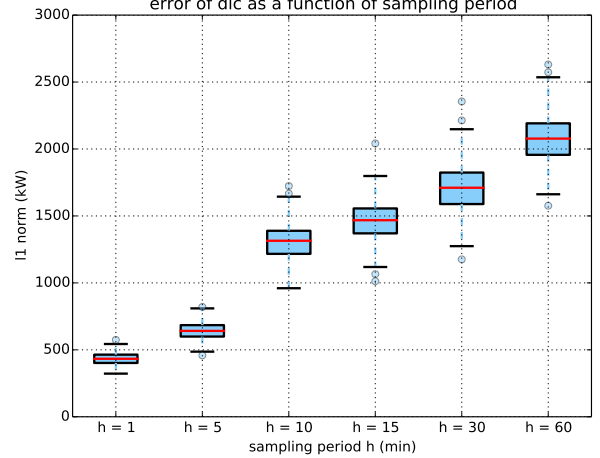


Fig. 3. A plot of how the error between the actual power consumed by the TCLs and the desired power consumption signal empirically varies with the sampling period h . The value we are plotting is $\|\sum_{i \in \mathcal{I}} P_i m_i - P_{\text{des}}\|_1$. The whiskers indicate all data points within 1.5 IQR. For reference, the error after 500 simulations of uncontrolled TCLs has an empirical mean of 5.39 MW with a standard error of 302 kW.

III. INFERENCEAL PRIVACY

In Section II, we consider how a larger sampling period affects the performance of a DLC program. Whereas it is intuitive that receiving fewer samples will increase a consumer's privacy, we can leverage results from nonintrusive load monitoring (NILM) to give guarantees of privacy to a user. In this section, we introduce our metric for privacy, called *inferential privacy*.

Data is collected from consumers with the intent of improving the efficiency of the smart grid. However, this data also contains information that does not assist in the operation of direct load control programs. This section quantifies how much information about the private lives of consumers is contained in data which is, in a sense, orthogonal to the information needed by the controller discussed in Section II.

A. A generative model of energy consumption

First, we present a generative model for a user's energy consumption patterns. We note that the aggregate power consumption signal is not private in and of itself; rather, it is the information about the consumer that can be inferred from this signal. Thus, suppose we have some private variable $\theta \in \Theta$, where Θ is some finite set. For example, θ could be a binary variable representing whether or not a user is home, or whether or not a high-power grow lamp is being used inside the building. We assume that for one fixed consumer, their type follows some multinomial distribution p_θ , i.e. $\theta \sim p_\theta$.

Given the consumer's private variable, a consumer will have a particular lifestyle, which will lead to different likelihoods of different device usage patterns. For example, a disengaged energy waster will likely leave her HVAC unit operating when she leaves her residence, whereas a green-advocate will likely turn her HVAC off whenever no one is present in her household. Thus, our model supposes a probability distribution a set of usage parameters, denoted by $u \in \mathcal{U}$, which is conditionally dependent on a user's type θ . Note here that we do not assume \mathcal{U} is finite. More formally, in our model we have $u \mid \theta \sim p_{u|\theta}(\cdot \mid \theta)$, where $p_{u|\theta}(\cdot \mid \theta)$ denotes the density of the distribution of these usage parameters conditioned on the consumer's private variable θ .

Finally, devices consume power based on how they are used. Given how a device is being used, its power consumption is independent of the consumer's private variable. Formally, this means that the power consumption $y \in \mathcal{Y}$ is based on the device usage, and is conditionally independent of the consumer's private variable given the device usage, i.e. $y \mid u, \theta \sim p_{y|u}(\cdot \mid u)$ where $p_{y|u}(\cdot \mid u)$ is the density of a probability distribution that models the devices inside the household. For example, in [29], [30], $p_{y|u}(\cdot \mid u)$ is the density of the output of a linear system with additive Gaussian noise. Additionally, note that \mathcal{Y} is not a finite set, and is usually taken to be \mathbb{R}^T , where T is the number of data samples collected.

To recap, our model for energy consumption in a household is a hierarchical Bayesian model where the customer's private variable is drawn according to some prior, the customer's device usage patterns is drawn from a distribution based on the customer's type, and the energy consumption of a customer is drawn from a distribution based on how devices are being used.

Assumption 1. *The consumer's private variable, device usage, and energy consumption are distributed according to the following hierarchical Bayes model:*

$$\theta \sim p_\theta \quad (3)$$

$$u \mid \theta \sim p_{u|\theta}(\cdot \mid \theta) \quad (4)$$

$$y \mid u, \theta \sim p_{y|u}(\cdot \mid u) \quad (5)$$

Here, θ takes values in a finite set Θ . Additionally, let $p_{y|\theta}(y \mid \theta) = \int p_{y|u}(y \mid u) p_{u|\theta}(u \mid \theta) du$ denote the density of $y \mid \theta$.

B. Connections to differential privacy

Recently, differential privacy has been a popular metric for measuring the privacy of users in a system [31], [32]. Differential privacy is a very attractive theoretical notion, as it is a broad definition which abstracts away the problem of defining an adversary model or privacy breach to provide a privacy guarantee independent of the adversary or definition of a privacy breach.

However, the concept of differential privacy often relies on additive noise of some form to give the privacy guarantees. In electrical grid applications, there are many cases where the original raw data is required, for practical, regulatory, performance, or economic reasons. To the best of our knowledge, there is no straightforward way to apply concepts of differential privacy to different sampling policies.

In contrast to differential privacy, our privacy metric can be used passively, i.e. without adding noise to the observed signal. It exploits the uncertainty intrinsic to device models and human behavior, i.e. the distributions on θ , u , and y . We also specify the definition of a privacy breach, i.e. an adversary correctly inferring a private parameter θ . Our definition is closer to the concept of equivocation metrics, e.g. [17], than differential privacy.

C. Inferential privacy

We suppose that a consumer's type θ is sensitive information: we wish to measure how much about θ is disclosed from y , the data transmitted by an AMI. Note that, if there are a finite number of ways to use devices in the household, i.e. \mathcal{U} is finite, device usage can be considered the sensitive information as a special case, by taking $\theta \equiv u$.

We note here that the problem of inferring u from y is a growing topic of research, known either as nonintrusive load monitoring or energy disaggregation [33], [34], [35], [36], [30]. People are actively working on recovering device usage information from an aggregate power consumption signal, and the privacy issue of inferring consumption patterns from AMI signals is a real threat.

Now, we introduce our adversary model.

Assumption 2. *Our adversary is able to observe the AMI signal y , and has knowledge of p_θ , $p_{u|\theta}$, and $p_{y|u}$. Additionally, this adversary has an arbitrary amount of computational power.*

This adversary has access to the measured data signal, and also holds priors on the consumer's private information θ . He also knows how different consumer types use devices, $p_{u|\theta}$, and also has access to models of the device's power consumption $p_{y|u}$. Although this adversary has quite a bit of knowledge about the consumers, he does not hold arbitrary side information.

Our privacy metric is the probability of error if an adversary tries to infer the private variable θ .

Definition 1. *Under the hierarchical Bayes model outlined in Assumption 1, an AMI protocol is 'α inferentially private' if, for any estimator $\hat{\theta} : \mathcal{Y} \rightarrow \Theta$, we have $\Pr(\hat{\theta}(y) \neq \theta) \geq \alpha$. This estimator can be based on information in p_θ , $p_{u|\theta}$, and $p_{y|u}$.*

Here, we note that this privacy metric is spread across Θ according to p_θ . As often arises in many statistical estimation problems, a privacy metric that guarantees privacy for every type is not a well-posed problem.

For example, suppose $\Theta = \{0, 1\}$, and consider the estimator $\hat{\theta} \equiv 0$. For any consumer of type $\theta = 0$, the adversary will correctly infer their type with this estimator. In other words, an adversary can always violate the privacy of one type of consumer by making the blanket assumption that everyone is a fixed type. In a sense, we gain privacy by noting that the adversary has to be successful across the different types Θ (weighted according to p_θ).

Regardless of the algorithm the adversary uses, we can bound the probability it will successfully breach a consumer's

privacy. Furthermore, this formula allows us to vary different parameters of the AMI, such as how often data is collected and transmitted. We will examine this on a concrete example in Section III-E. This guarantee is also simple for consumers to interpret, and can be used in the design of privacy contracts between the utility company and electricity consumers [37].

We note that it may not be realistic to suppose the adversary has access to this information. However, any adversary who tries to infer θ from y with less information will only do worse than our adversary model. Thus, this model provides a conservative estimate against all weaker adversary models.

D. Theory

In this section, we will derive results that allow us to calculate values of α that satisfy the condition given in Definition 1. This section is an extension of our work [38].

There are three methods by which we can derive lower bounds. Depending on the particular form of p_θ , $p_{u|\theta}$, and $p_{y|u}$, some forms of the lower bound may be easier to calculate than others.

1) *Likelihood-based methods*: Let $\Theta = \{1, \dots, r\}$. We can define the maximum a posteriori (MAP) estimator $\hat{\theta}_{MAP}$, which maximizes $\Pr(\hat{\theta}(y) = \theta)$.

Proposition 1. [39], [38] *Under the hierarchical Bayes model outlined in Assumption 1, $\Pr(\hat{\theta}(y) = \theta)$ is maximized by:*

$$\hat{\theta}_{MAP}(y) = \arg \max_i (\pi(i) \cdot p_{y|\theta}(y|i)) \quad (6)$$

The optimality of the MAP estimator with respect to the prior π immediately leads to a guarantee of privacy.

Proposition 2. [38] *Under the hierarchical Bayes model outlined in Assumption 1, the AMI protocol is α inferentially private, where $\alpha = \Pr(\hat{\theta}_{MAP}(y) \neq \theta)$. Furthermore, the AMI protocol is not α' inferentially private for any $\alpha' > \alpha$.*

Although this bound is optimal, it is often difficult to calculate. In these instances, some of the latter bounds may be used as a surrogate.

2) *Le Cam's method*: We present Le Cam's lemma in the context of our energy consumption model. Again, let $\Theta = \{1, 2, \dots, r\}$. First, we offer two definitions of distances between probability distributions.

Definition 2. *The total variation distance between two densities p and q on a measure space (X, \mathcal{A}, μ) is given by:*

$$\|p - q\|_{TV} = \sup_{A \in \mathcal{A}} \left| \int_A p(x) - q(x) \mu(dx) \right| = \quad (7)$$

$$\frac{1}{2} \int_X |p(x) - q(x)| \mu(dx) \quad (8)$$

Definition 3. *The Kullback-Leibler (KL) divergence between two densities p and q on a measure space (X, \mathcal{A}, μ) is given by:*

$$D_{KL}(p\|q) = \int p(x) \log \frac{p(x)}{q(x)} \mu(dx) \quad (9)$$

Similarly, we will define the KL divergence between two random variables X and Y to be the KL divergence between their densities, and it will be denoted $D_{KL}(X\|Y)$.

Now, we can state Le Cam's lemma.

Proposition 3. [40], [41] *Assume the hierarchical Bayes model outlined in Assumption 1. Then, for any estimator $\hat{\theta} : \mathcal{Y} \rightarrow \Theta$ and any distinct $i, j \in \Theta$, we have:*

$$\Pr(\hat{\theta}(y) \neq \theta \mid \theta = i) + \Pr(\hat{\theta}(y) \neq \theta \mid \theta = j) \geq \quad (10)$$

$$1 - \|p_{y|\theta}(\cdot|i) - p_{y|\theta}(\cdot|j)\|_{TV} \quad (11)$$

A quick corollary is a lower bound on the probability of error:

Proposition 4. *Under the assumptions of Proposition 6, $\Pr(\hat{\theta} \neq \theta)$ is bounded below by:*

$$\min(\pi(i), \pi(j)) \cdot (1 - \|p_{y|\theta}(\cdot|i) - p_{y|\theta}(\cdot|j)\|_{TV}) \quad (12)$$

In practice, it will suffice to find an over-approximation of the total variation distance. For example, we have Pinsker's inequality:

Proposition 5. [42] *For any densities p and q :*

$$\|p - q\|_{TV} \leq \sqrt{\frac{1}{2} D_{KL}(p\|q)} \quad (13)$$

Thus, we can provide a guarantee of inferential privacy.

Proposition 6. *Under the hierarchical Bayes model outlined in Assumption 1, the AMI protocol is α inferentially private, where:*

$$\alpha = \max_{i \neq j} [\min(\pi(i), \pi(j)) \cdot (1 - \|p_{y|\theta}(\cdot|i) - p_{y|\theta}(\cdot|j)\|_{TV})] \quad (14)$$

3) *Fano's method*: Here we will state Fano's inequality in the context of our energy consumption model, where $\Theta = \{1, \dots, r\}$.

Proposition 7. [41] *In the model of Assumption 1, for any estimator $\hat{\theta} : \mathcal{Y} \rightarrow \Theta$, the probability of error $P(\hat{\theta}(y) \neq \theta)$ is bounded below by:*

$$\frac{1}{\log(r-1)} \left[\log r - \frac{1}{r^2} \sum_{i,j} D_{KL} [p_{y|\theta}(\cdot|i) \| p_{y|\theta}(\cdot|j)] - \log 2 \right] \quad (15)$$

Thus, we have the quick corollary:

Proposition 8. *Under the hierarchical Bayes model outlined in Assumption 1, the AMI protocol is α inferentially private, where α is given by Equation 15.*

E. Privacy metric example

We instantiate our privacy metric on a concrete example here. Consider the following scenario. Households consider their income private. However, their income levels will affect their behaviors at home; in this paper, we focus on how their cooking behaviors change. To model this, we use data from the U.S. Energy Information Administration's 2009 Residential Energy Consumption Survey (RECS) [43]. By observing these different cooking behaviors through a household's energy consumption, an adversary may infer the income of the household.

Formally, let $\Theta = \{\theta_L, \theta_M, \theta_U\}$ denote the private parameter corresponding to lower (less than \$20,000), middle (\$20,000 to \$59,999), and upper (\$60,000 or more) class incomes. Across 113.6 million U.S. homes, 23.7 million households are θ_L , 48.7 million are θ_M , and 41.2 million are θ_U [43]. This will be our prior, p_θ .

Furthermore, we look at the overall energy consumption of each consumer type. This data is shown in Figure 4. For each type, we fit a log-normal distribution to the overall energy consumption. To estimate the location parameter μ and scale parameter σ , we used the unbiased, minimum variance estimators [44, Chapter 4] on the log of the data³. We assume the scale parameter is the same for all three private parameters, and we can see that these distributions approximate the data quite well.

Thus in this framework, θ determines u , which in this example, is μ , the location parameter. In other words, $p_{u|\theta}$ has a trivial distribution for each value of θ . This, in turn, determines the distribution across the observable energy consumption y , i.e. $p_{y|u}(\cdot | u)$ is the density of a $\ln N(\mu, \sigma^2)$ distribution.

We assume that power consumption on smaller time scales is distributed similarly to this annual data, and these distributions are independent across time.

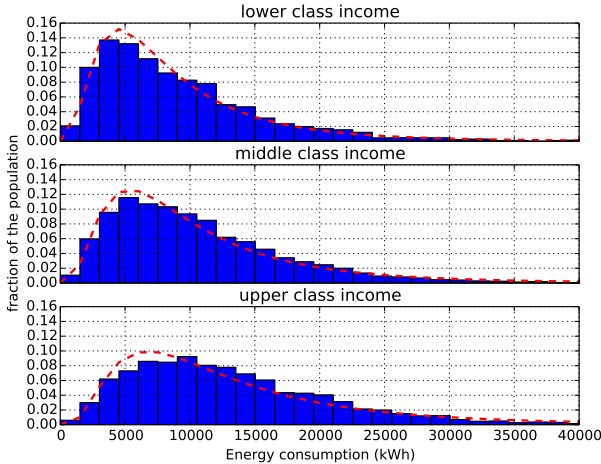


Fig. 4. Histograms of the United States household total annual energy consumptions in each income level in 2009 [43], corresponding to private parameters θ_L , θ_M , and θ_U . The data roughly follows a log-normal distribution. The location parameters are $\mu_L = 8.88$, $\mu_M = 9.06$, and $\mu_H = 9.31$, and we assumed all three distributions had the sample scale parameter, $\sigma = 0.49$. To model sampling, we assume that this data is representative of energy consumption on smaller time scales as well.

With this assumption, we can consider the distribution of energy consumption at different sampling rates. Note that, if we sample at high frequencies, we receive more measurements than in the low frequency case, but each measurement is less

informative with regards to the consumer's income level⁴.

Since the scale parameters are the same for all 3 distributions, we can explicitly calculate the MAP using the theory of exponential families [44, Chapter 2]. Then, using Proposition 2, we can calculate the probability an adversary can infer the private parameters, i.e. income level, from the AMI signals. This is represented in Figure 5.

We can see that very high frequency data provides little guarantees of privacy of income level, but this privacy level, α , quickly increases as the sampling period h increases. This is likely due to the fact that the number of measurements used to calculate the MAP decreases quickly for these values of h . For example, if $h = 1$, then 60 samples are used to calculate the MAP; if $h = 10$, then only 6 samples are used.

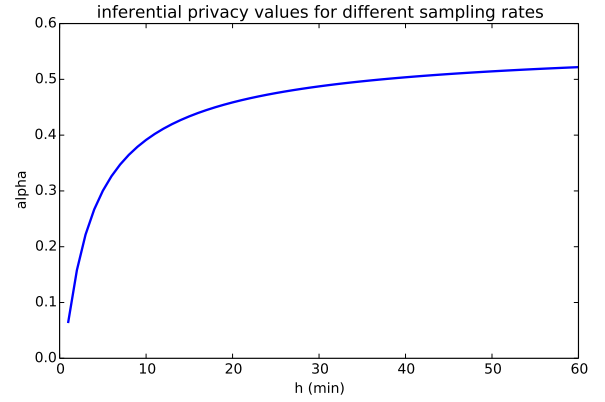


Fig. 5. A plot of the inferential privacy value α , as a function of the sampling period h . This is the privacy value for measurements in one hour; we would expect this to decrease if we sampled for longer durations. Note that our framework accounts for the fact that with longer sampling periods, we receive fewer measurements, but each individual measurement is more informative.

Although we focus on a particular example here, this framework can be applied to more detailed models, i.e. more informed adversaries, and other private parameters. For example, we consider the case where an adversary has knowledge of correlation across time and high frequency dynamics across time in [38].

IV. CONCLUSIONS AND FUTURE WORK

In an effort to quantify the tradeoff between smart grid operations and adversarial inferences about consumer behavior, we consider a direct load control of thermostatically controlled loads, and analyze how its performance degrades as it receives samples less and less frequently—a privacy preserving metering policy. We introduce a new privacy metric, *inferential privacy*, that exploits the uncertainty intrinsic to device models and human behavior. Our contribution is a framework for understanding the utility of data in direct

³Recall that the log-normal distribution, denoted $\ln N(\mu, \sigma)$, is defined by a location parameter μ and scale parameter σ , with density $x \mapsto \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$ for $x > 0$.

⁴Here, we scale the data according to the time scale, and, as before, we used the uniform, minimum variance estimators on the log of the data [44, Chapter 4]. For example, if we receive measurements every minute, the location parameters for each measurement are $\mu_L = 0.014$, $\mu_M = 0.016$, and $\mu_H = 0.017$, whereas if we receive measurements hourly, the location parameters are $\mu_L = 0.82$, $\mu_M = 0.99$, and $\mu_H = 1.26$.

load control programs, as well as understanding the private information about consumers contained in the data.

The introduction of this new privacy metric leads to a number of interesting open questions. For instance, we could consider that consumers' private parameters live on a continuum or are time-varying as is the case in practice. Another interesting direction is in quantifying the network effects of privacy. In particular, what information from the crowd to infer private information from the individual and how does this scale as the number of users increases? In terms of understanding the impact on smart grid operations, such privacy metrics can be used in conjunction with economic tools that capture individual preferences over energy consumption and privacy to help balance this utility-privacy tradeoff.

REFERENCES

- [1] M. Lisovich, D. Mulligan, and S. Wicker, "Inferring personal information from demand-response systems," *IEEE Security Privacy*, vol. 8, pp. 11–20, Jan 2010.
- [2] R. Anderson and S. Fuloria, "On the security economics of electricity metering," Ninth Workshop on the Economics of Information, 2010.
- [3] G. Smith, "Marijuana bust shines light on utilities," Post and Courier, Jan 2012.
- [4] Government Accountability Office, "Electricity grid modernization: Progress being made on cybersecurity guidelines, but key challenges remain to be addressed," Jan 2011.
- [5] K. Kursawe, G. Danezis, and M. Kohlweiss, "Privacy-friendly aggregation for the smart-grid," in *Proc. of the 11th Int. Conf. on Privacy Enhancing Technologies*, ser. PETS'11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 175–191.
- [6] A. Rial and G. Danezis, "Privacy-preserving smart metering," in *Proc. of the 10th Annu. ACM Workshop on Privacy in the Electronic Society*, ser. WPES '11. New York, NY, USA: ACM, 2011, pp. 49–60.
- [7] G. Taban and V. Gligor, "Privacy-preserving integrity-assured data aggregation in sensor networks," in *Int. Conf. on Computational Science and Engineering*, vol. 3, Aug 2009, pp. 168–175.
- [8] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," in *2010 1st IEEE Int. Conf. on Smart Grid Communications (SmartGridComm)*, Oct 2010, pp. 327–332.
- [9] G. Acs and C. Castelluccia, "I Have a DREAM! (Differentially private smArt Metering)," in *Information Hiding*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2011, vol. 6958, pp. 118–132.
- [10] The Smart Grid Interoperability Panel - Cyber Security Working Group, "NISTIR7628 – Guidelines for Smart Grid Cyber Security: Vol. 2, Privacy and the Smart Grid."
- [11] North American Energy Standards Board, "NAESB privacy policy."
- [12] Department of Energy, "Data access and privacy issues related to smart grid technologies."
- [13] Public Utility Commission of Texas, "Electric Substantive Rules – Chapter 25."
- [14] California Public Utilities Commission, "Decision adopting rules to protect the privacy and security of the electricity usage data of the customers of Pacific Gas and Electric Company, Southern California Edison Company, and San Diego Gas & Electric Company."
- [15] A. A. Cárdenas, S. Amin, G. Schwartz, R. Dong, and S. S. Sastry, "A game theory model for electricity theft detection and privacy-aware control in AMI systems," in *Proc. of the 50th Allerton Conf. on Communication, Control, and Computing*, 2012, pp. 1830–1837.
- [16] L. Sankar, S. Kar, R. Tandon, and H. Poor, "Competitive privacy in the smart grid: An information-theoretic approach," in *2011 IEEE Int. Conf. on Smart Grid Communications (SmartGridComm)*, Oct 2011, pp. 220–225.
- [17] L. Sankar, S. Rajagopalan, and H. Poor, "Utility-privacy tradeoffs in databases: An information-theoretic approach," *IEEE Trans. Inf. Forens. Security*, vol. 8, pp. 838–852, June 2013.
- [18] D. S. Callaway, "Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy," *Energy Conversion and Management*, vol. 50, no. 5, pp. 1389 – 1400, 2009.
- [19] C. Perfumo, E. Kofman, J. H. Braslavsky, and J. K. Ward, "Load management: Model-based control of aggregate power for populations of thermostatically controlled loads," *Energy Conversion and Management*, vol. 55, pp. 36 – 48, 2012.
- [20] M. Alexander, K. Agnew, and M. Goldberg, "New approaches to residential direct load control in California," in *2008 ACEEE Summer Study on Energy Efficiency in Buildings*, 2008.
- [21] California Energy Commission, "Docket No. 13-IEP-1F: Increasing demand response capabilities in California," 2013.
- [22] N. Lu, "An evaluation of the hvac load potential for providing load balancing service," *IEEE Trans. on Smart Grid*, vol. 3, no. 3, pp. 1263–1270, Sept 2012.
- [23] N. Lu and Y. Zhang, "Design considerations of a centralized load controller using thermostatically controlled appliances for continuous regulation reserves," *IEEE Trans. on Smart Grid*, vol. 4, no. 2, pp. 914–921, June 2013.
- [24] J. Mathieu, S. Koch, and D. Callaway, "State estimation and control of electric loads to manage real-time energy imbalance," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 430–440, 2013.
- [25] N. Ruiz, I. Cobelo, and J. Oyarzabal, "A direct load control model for virtual power plant management," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 959–966, May 2009.
- [26] S. Moura, J. Bendtsen, and V. Ruiz, "Observer design for boundary coupled pdes: Application to thermostatically controlled loads in smart grids," in *IEEE 52nd Annu. Conf. on Decision and Control (CDC)*, Dec 2013, pp. 6286–6291.
- [27] D. Callaway and I. Hiskens, "Achieving controllability of electric loads," *Proc. of the IEEE*, vol. 99, no. 1, pp. 184–199, Jan 2011.
- [28] California Independent System Operators, "Business practice manual for market operations," January 2014.
- [29] R. Dong, L. Ratliff, H. Ohlsson, and S. Sastry, "A dynamical systems approach to energy disaggregation," in *2013 IEEE 52nd Annu. Conf. on Decision and Control (CDC)*, Dec 2013, pp. 6335–6340.
- [30] R. Dong, L. J. Ratliff, H. Ohlsson, and S. S. Sastry, "Energy disaggregation via adaptive filtering," in *2013 51st Annu. Allerton Conf. on Communication, Control, and Computing (Allerton)*, Oct 2013, pp. 173–180.
- [31] C. Dwork, "Differential privacy," in *Proc. of the Int. Colloq. on Automata, Languages and Programming*. Springer, 2006, pp. 1–12.
- [32] J. Le Ny and G. Pappas, "Differentially private filtering," *IEEE Trans. Autom. Control*, vol. 59, pp. 341–354, Feb 2014.
- [33] G. Hart, "Nonintrusive appliance load monitoring," *Proc. of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.
- [34] S. Gupta, M. S. Reynolds, and S. N. Patel, "Electrisense: single-point sensing using EMI for electrical event detection and classification in the home," in *Proc. of the 12th ACM Int. Conf. on Ubiquitous Computing*, ser. Ubicomp '10. New York, NY, USA: ACM, 2010, pp. 139–148.
- [35] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial HMMs with application to energy disaggregation," in *Proc. of the Int. Conf. on Artificial Intelligence and Statistics*, 2012.
- [36] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Nonintrusive load monitoring using prior models of general appliance types," in *Proc. of the 26th AAAI Conf. on Artificial Intelligence*, 2012.
- [37] L. J. Ratliff, C. Barreto, R. Dong, H. Ohlsson, A. Cárdenas, and S. S. Sastry, "Effects of risk on privacy contracts for demand-side management," *arXiv*, 2015.
- [38] R. Dong, L. Ratliff, H. Ohlsson, and S. S. Sastry, "Fundamental limits of nonintrusive load monitoring," in *Proc. of the 3rd Int. Conf. on High Confidence Networked Systems*, ser. HiCoNS '14. New York, NY, USA: ACM, 2014, pp. 11–18. [Online]. Available: <http://doi.acm.org/10.1145/2566468.2566471>
- [39] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.
- [40] L. Le Cam, "Convergence of estimates under dimensionality restrictions," *The Ann. of Statistics*, vol. 1, no. 1, pp. pp. 38–53, 1973. [Online]. Available: <http://www.jstor.org/stable/2958155>
- [41] B. Yu, "Assouad, Fano, and Le Cam," in *Festschrift for Lucien Le Cam*. Springer, 1997, pp. 423–435.
- [42] A. B. Tsybakov, *Introduction to Nonparametric Estimation*. Springer New York, 2009.
- [43] C. Berry, "Residential energy consumption survey," U.S. Energy Information Administration, Tech. Rep., 2009.
- [44] R. W. Keener, *Theoretical Statistics: Topics for a Core Course*. Springer, 2010.